AMERIFLUX
MANAGEMENT PROJECT

Data QA/QC
Self-Review & Automation

Housen Chu, Danielle Svehla Christianson
Deb Agarwal, You-Wei Cheah, Rachel Hollowgrass

AmeriFlux Online Data Workshop, February 23/24, 2022

Workshop Goals

Who should attend?

- Site teams submit flux/met data regularly
- Site teams have interest in participating in the development and test

Presentation (~50 minutes)

- Data pipeline overview
- QA/QC process and test modules
- Q & A (~10 minutes)
- Plan for this year and next stage

Discussion (~40 minutes)

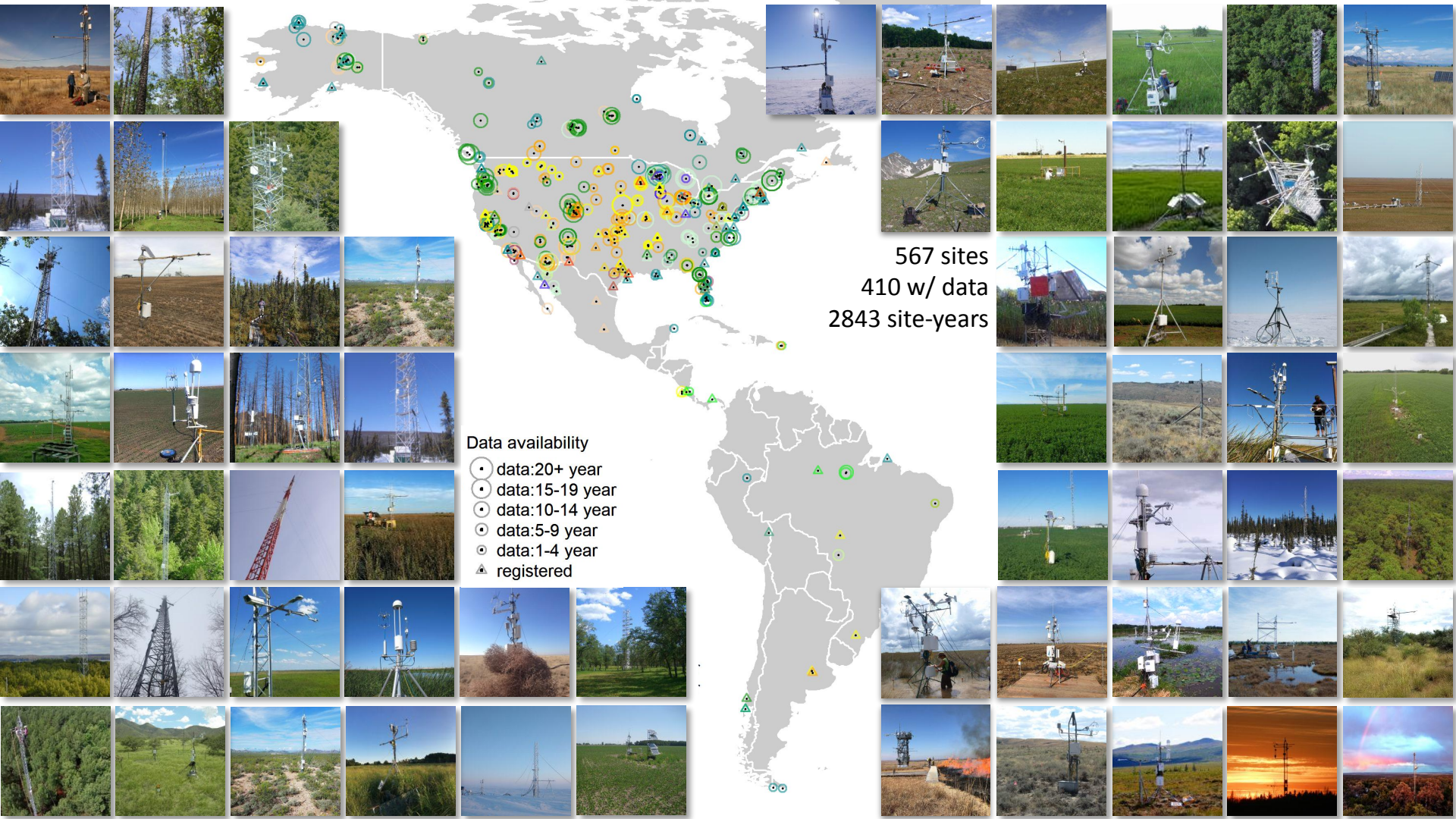
- https://docs.google.com/document/d/1HuDNhYrXUMGB0YDMNcZjkP1Yr2GjsbHVprpk-HiBjxU/edit?usp=sharing

The presentation will be recorded and made available on AmeriFlux website.

Mute during presentation.
Unmute for Q & A.

View chat window.
Send messages for questions, comments & zoom help.

Technical support in webinar
AMP-webinars@lbl.gov


567 sites
410 w/ data
2843 site-years
Data availability
data:20+ year
data:15-19 year
data:10-14 year
data:5-9 year
data:1-4 year
registered

AmeriFlux Flux/Met data Pipeline

Site Teams

AmeriFlux Management Project (AMP)

Available to Data Users

AmeriFlux Flux/Met data Pipeline

Site Teams

AmeriFlux Management Project (AMP)

BASE Data Product

- Provided as submitted by site team
- QA/QC on general data quality
- All FP variables supported
- All levels of aggregation supported

Available to Data Users

FLUXNET Data Product

- Value-added data product
- QA/QC on ONEFlux requirements
- Subset of standard FP variables
- Site-representative aggregation

AmeriFlux Flux/Met data Pipeline

Site Teams

Submit Flux/Met data

AmeriFlux Management Project (AMP)

Format QA/QC → Data QA/QC → BASE Publish → ONEFlux Processing → FLUXNET Publish

BASE Data Product

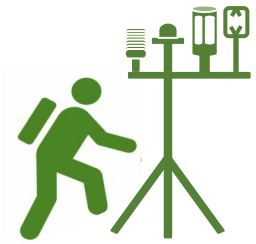
- Provided as submitted by site team
- QA/QC on general data quality
- All FP variables supported
- All levels of aggregation supported

Available to Data Users

FLUXNET Data Product

- Value-added data product
- QA/QC on ONEFlux requirements
- Subset of standard FP variables
- Site-representative aggregation

AmeriFlux Flux/Met data Pipeline



Site Teams

Submit Flux/Met data

~ 1-6 months

~ 1-2 days

Iteration to verify or fix issues

per upload
~ Immediately

batch
~ 1-2 months

batch
~ 1-2 months

Format QA/QC

Data QA/QC

BASE Publish

ONEFlux Processing

FLUXNET Publish

AmeriFlux Management Project (AMP)

BASE Data Product

- Provided as submitted by site team
- QA/QC on general data quality
- All FP variables supported
- All levels of aggregation supported

Available to Data Users

FLUXNET Data Product

- Value-added data product
- QA/QC on ONEFlux requirements
- Subset of standard FP variables
- Site-representative aggregation

AmeriFlux Flux/Met data Pipeline

Site Teams

Submit Flux/Met data

~ 1-6 months

~ 1-2 days

Iteration to verify or fix issues

per upload
~ Immediately

batch
~ 1-2 months

batch
~ 1-2 months

Format QA/QC → Data QA/QC → BASE Publish → ONEFlux Processing → FLUXNET Publish

AmeriFlux Management Project (AMP)

Recording available:

- Post-submission data life cycle: FP-In to BASE publishing
- Requirements for processing a site using ONEFlux
- https://ameriflux.lbl.gov/community/amp-webinar-series/

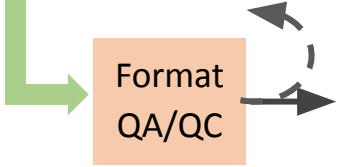
BASE Data Product

- Provided as submitted by site team
- QA/QC on general data quality
- All FP variables supported
- All levels of aggregation supported

Available to Data Users

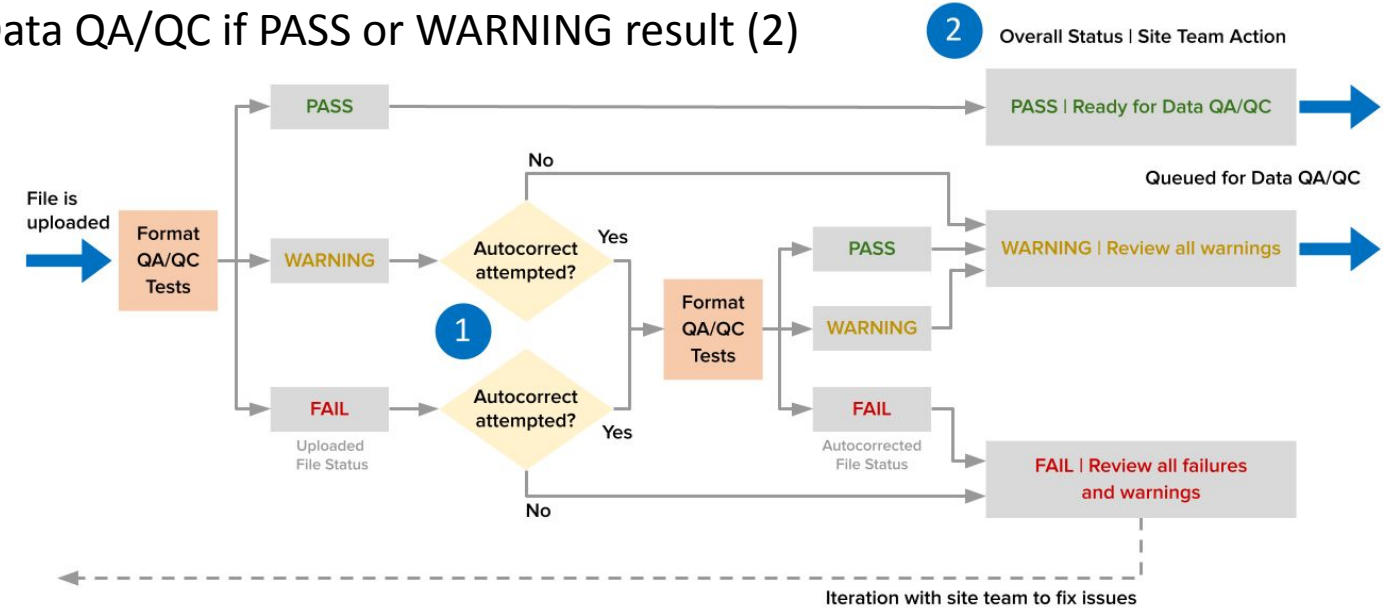
FLUXNET Data Product

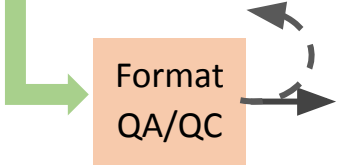
- Value-added data product
- QA/QC on ONEFlux requirements
- Subset of standard FP variables
- Site-representative aggregation



Format QA/QC Processing

- Fully-automated process begins immediately upon upload
- Format QA/QC Report links emailed to uploader ~15 minutes after upload
- Per file evaluation
- One attempt to auto-correct any issues (1)
- File is queued for Data QA/QC if PASS or WARNING result (2)





Format QA/QC Report Email

AmeriFlux QAQC-3065 Format Results - ACTION REQUIRED | CC-sss data uploaded on Sep 04, 2019 Inbox ×

Dear Danielle Christianson,

Thank you for uploading data for CC-sss on Sep 04, 2019.

Format QA/QC results

CC-sss_HR_200001011000_200001012000_bad29.csv:

- **FAIL | Replacement file required.**
- Read details in this report: https://ameriflux.lbl.gov/qaqc-report/?site_id=CC-sss&report_id=63097

CC-sss_HR_200001011000_200001012000_scinot.csv:

- **FAIL | Replacement file required.**
- Read details in this report: https://ameriflux.lbl.gov/qaqc-report/?site_id=CC-sss&report_id=63096

Format QA/QC assesses the compliance of your data submission with AmeriFlux FP-In format (https://ameriflux.lbl.gov/half-hourly-hourly-data-upload-format/. If needed, you can re-upload your data at https://ameriflux.lbl.gov/data/upload-data/ and/or reply to this email to discuss with us.

View the status of all your uploaded files at https://ameriflux.lbl.gov/qaqc-reports-data-team/.

If all files passed Format QA/QC and there are no pending issues for your site, Data QA/QC will be run. You can track communications on this Format QA/QC report at QAQC-3065 using your AmeriFlux account ID and password to login.

Sincerely,
AMP Data Team

Format QA/QC

Format QA/QC Report

QA/QC Report: Format

This report details results of the AmeriFlux QA/QC data processing pipeline.
For more information, see How to Read This Report, QA/QC Results Definitions,

PASS **Ready for Data QAQC**
No further action needed by the site team.

Uploaded File Report US-PFa_HR_201801010000_201901010000.csv

Site ID: US-PFa
Site contact: Ankur Desai

Uploader: Ankur Desai
Upload date: 2018-Jul-16 11:44
Uploaded filename: US-PFa_HR_201801010000_201901010000-20180

Format QA/QC report summary:
All format QA/QC tests attempted. No issues were encountered. AMP will

Test	Results	Ad
All Format QA/QC tests passed.	✔ PASS	

Variable names found in the file:
TIMESTAMP_START, TIMESTAMP_END, CO2_1_1_1, CO2_1_2_1, CO2
CH4_1_1_1, CH4_1_2_1, CH4_1_3_1, FC_1_1_1, FC_1_2_1, FC_1_3_1
SCH4_1_1_1, H, H_1_1_1, H_1_2_1, H_1_3_1, LE, LE_1_1_1, LE_1_2_
SLE_1_2_1, SLE_1_3_1, WD_1_1_1, WD_1_2_1, WD_1_3_1, WD_F_1_
USTAR_1_1_1, USTAR_1_2_1, USTAR_1_3_1, USTAR_F_1_3_1, PA_1_
VPD_F_1_3_1, SWC_1_1_1, PPFD_IN_1_1_1, P, NEE, NEE_F, NEE_1_

Processing code version: 0.4.19
Processing log file: http://ameriflux-data.lbl.gov/QAQCLogs/QAQC_repo

QA/QC Report: Format

This report details results of the AmeriFlux QA/QC data processing pipeline.
For more information, see How to Read This Report, QA/QC Results Definitions, FAQ, and Up

WARNING **Review all warnings**
If autocorrected file is OK, no action is needed by the site team. If correcti

Autocorrected File Report US-MOz_HH_200501010000_200601010000.csv

Site ID: US-MOz
Site contact: Jeffrey Wood

Uploader: AMP Data Team (original file uploaded by Format QAQC Pipeline)
Upload date: 2018-Aug-15 17:27
Uploaded filename: US-MOz_HH_200501010000_200601010000-20180815172726

Format QA/QC report summary:
All format QA/QC tests attempted. Issues were encountered. AMP attempted to auton
warnings below. If autocorrected file is OK, no action is needed by the site team. If cor

Test	Results	Additional Info
AMP made these autocorrections.	? WARNING	• Filename cor time)
Any Variables suspected gap-fill?	? WARNING	These variables no missing valu
Any Variables with ALL Data Missing?	? WARNING	These variables **H_1_1_1**. Previ be overwritten.

Variable names found in the file:
TIMESTAMP_START, TIMESTAMP_END, P_1_1_1, PPFD_IN_1_1_1, PPFD_OUT_1
LW_IN_1_1_1, LW_OUT_1_1_1, NETRAD_1_1_1, TA_1_1_1, RH_1_1_1, CO2_1_1_
USTAR_1_1_1, TS_1_1_1, SWC_1_1_1, G_1_1_1, PA_1_1_1, FC_1_1_1, SC_1_1_
NEE, NEE_F

Processing code version: 0.4.23
Processing log file: http://ameriflux-data.lbl.gov/QAQCLogs/QAQC_report_US-MOz_

Uploaded File Report US-MOz_HH_200501010000000_200601010000000.csv

⚠ **Consider** revising your file preparation for future submissions by opening and reviewing t

QA/QC Report: Format

This report details results of the AmeriFlux QA/QC data processing pipeline.
For more information, see How to Read This Report, QA/QC Results Definitions, FAQ, and Upload Format Instructions

FAIL **Review failures and warnings**
Upload a corrected replacement file.

Uploaded File Report Tonzi-understory-2016.dat Report ID: 8052

Site ID: US-Ton
Site contact: Dennis Baldocchi

Uploader: Siyan Ma
Upload date: 2017-Mar-22 14:26
Uploaded filename: Tonzi-understory-2016-2017011912065058.dat

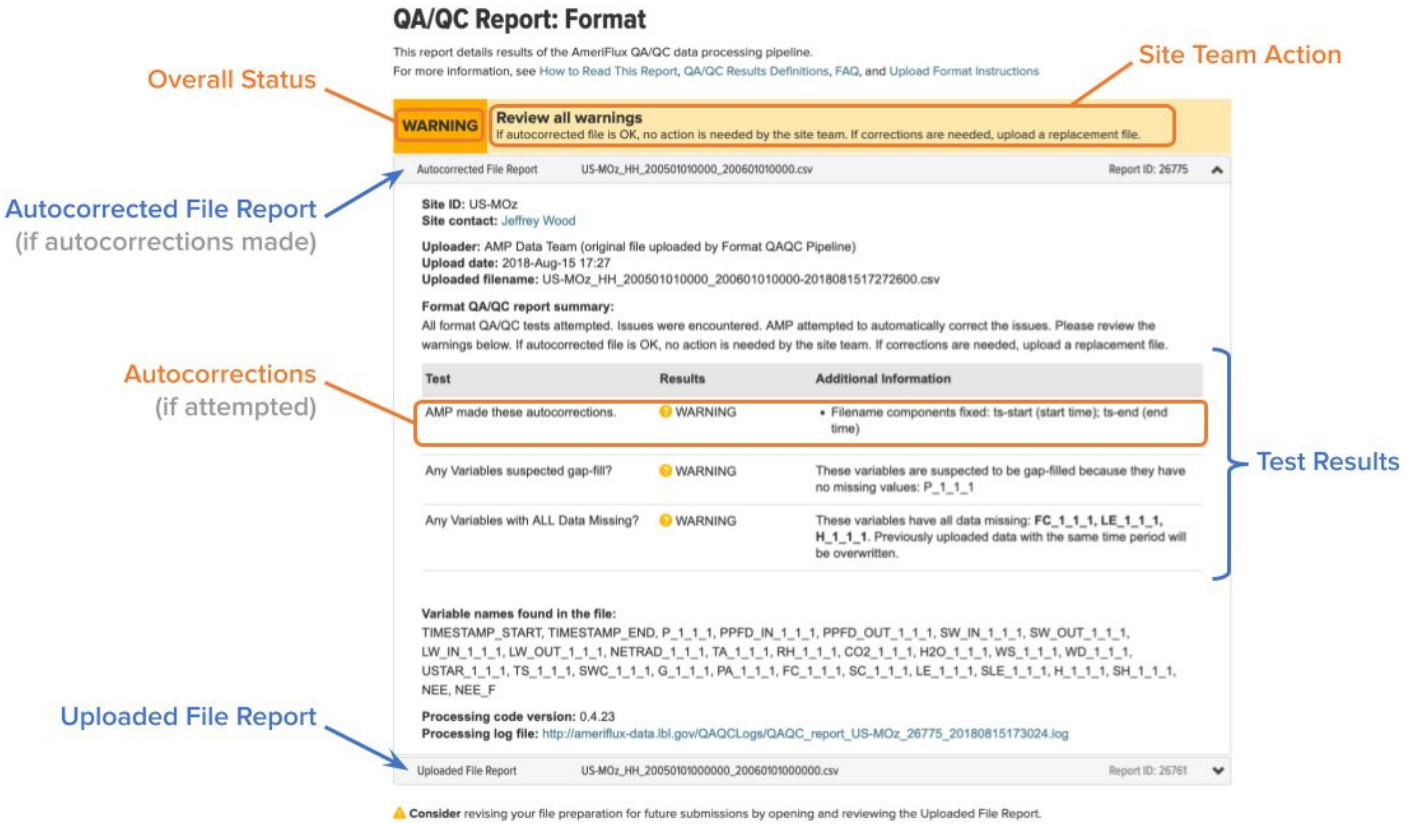
Format QA/QC report summary:
All format QA/QC tests attempted. Issues were encountered. Please correct issues and upload a replacement file.

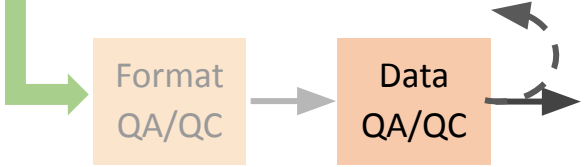
Test	Results	Additional Information
Are Timestamp variables present?	⚠ FAIL	Expected timestamp variable(s) **TIMESTAMP_START, TIMESTAMP_END** is / are missing.
Timestamp problem encountered.	⚠ FAIL	Filename Matches File Contents, Timestamp Column Resolution, Timestamp Row Resolution, Timestamp Duplicates
Issues that cannot be autocorrected.	⚠ FAIL	Unable to repair timestamps. AutoRepair FAILED.
Is Filename Format valid?	⚠ FAIL	These filename components are not in the standard AmeriFlux format: **extension is not csv**
Are Timestamp variables as expected?	⚠ FAIL	These unexpected variables were found in columns 1 & 2 instead of TIMESTAMP_START and TIMESTAMP_END: **yr, day**
Is Filename Format valid?	? WARNING	These filename components are not in the standard AmeriFlux format: **incorrect number of components (expect timestamp errors)**
Are Data Variable names in correct format?	? WARNING	These variable names are not in standard AmeriFlux format: **yr, day, endhour, endmin, DOY, FC_WPL_2D, fc_flag, WC_2D, CO2_LI7500, RHOC, CO2_var, CO2_skewness, CO2_kurtosis, RHOQ, q_var, q_skewness, q_kurtosis, Tsonic, Tsonic_var, Tsonic_skewness, Tsonic_kurtosis, Wind_Direction, Wind_Velocity, Friction_Velocity, stdw, wbar, w_var, w_kurtosis, u2D_var, v2D_var, Tair, absolute_humidity, Vapor_pressure_deficit, Relhumidity, Pressure, TSOIL2, TSOIL4, TSOIL8, TSOIL16, TSOIL32, soil_moisture_00cm, soil_moisture_20cm, soil_moisture_50cm, precipitation**. They will not be included in the standard AmeriFlux data products.

Format QA/QC

Format QA/QC Report

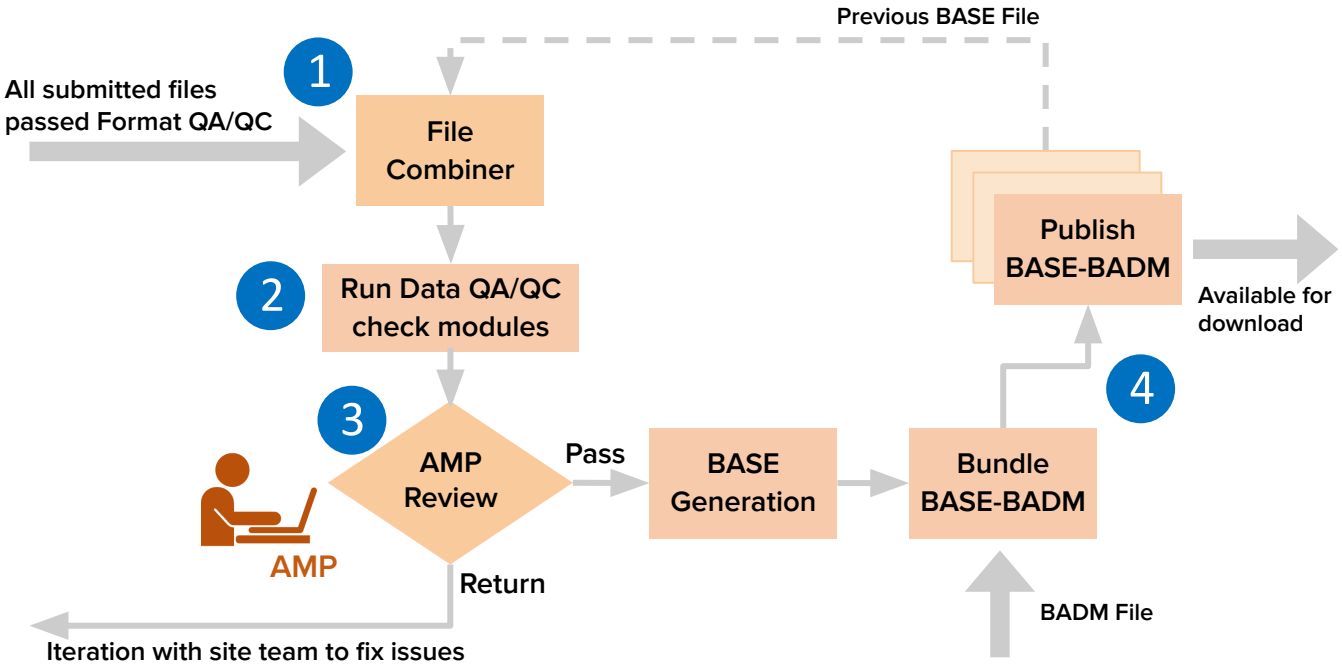
New online QA/QC Documentation available
https://ameriflux.lbl.gov/data/data-processing-pipelines/format-qaqc/

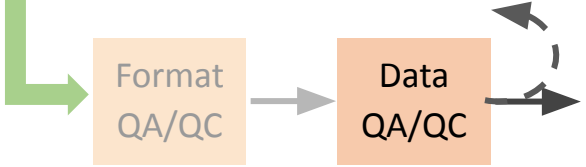




Data QA/QC and Report Email

- Most recently uploaded data that pass Format QA/QC
- Performed on entire data record (recent uploaded + previous BASE)
- No data filtering/correction done by AMP

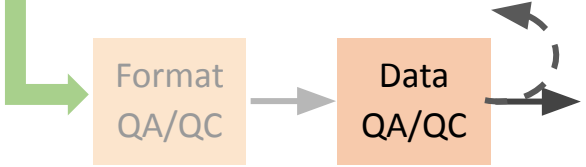




Data QA/QC Modules

- Secondary independent check
  - Adapted from Pastorello et al., 2014, IEEE, DOI: 10.1109/eScience.2014.45
  - Lesson-learned from flux networks
  - Feedback from data users and ONEFlux processing
  - Post hoc approach (w/ limited ancillary info & diagnostics)
  - Trade-off between site-specific & universal
  - Emphasize visualization (for issue identification & communication)
- Adopted in FLUXNET2015 & all BASE published since 2017

https://ameriflux.lbl.gov/data/data-processing-pipelines/data-qaqc/



Data QA/QC Modules

- Secondary independent check
  - Adapted from Pastorello et al., 2014, IEEE, DOI: 10.1109/eScience.2014.45
  - Lesson-learned from flux networks
  - Feedback from data users and ONEFlux processing
  - Post hoc approach (w/ limited ancillary info & diagnostics)
  - Trade-off between site-specific & universal
  - Emphasize visualization (for issue identification & communication)
- Adopted in FLUXNET2015 & all BASE published since 2017

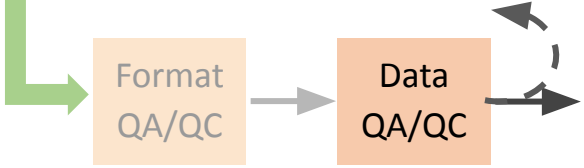
Implemented Modules

- Timestamp alignment
- Physical range
- Diurnal & Seasonal pattern
- Multivariate comparison
- USTAR filtering

Planned Modules

- Variable availability
- Sign convention check
- SIGMA family
- Variability check
- Unit check (ratio-percentage, metric conversion)

https://ameriflux.lbl.gov/data/data-processing-pipelines/data-qaqc/



Check Module

5 Target Variables

3 *Supporting Summary Statistics*

4
- Identified Issues
- ……

1

Figure

2

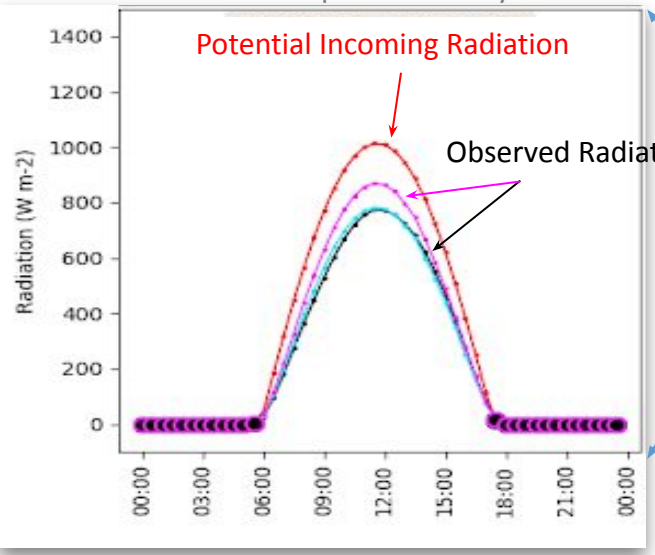
Additional Figure

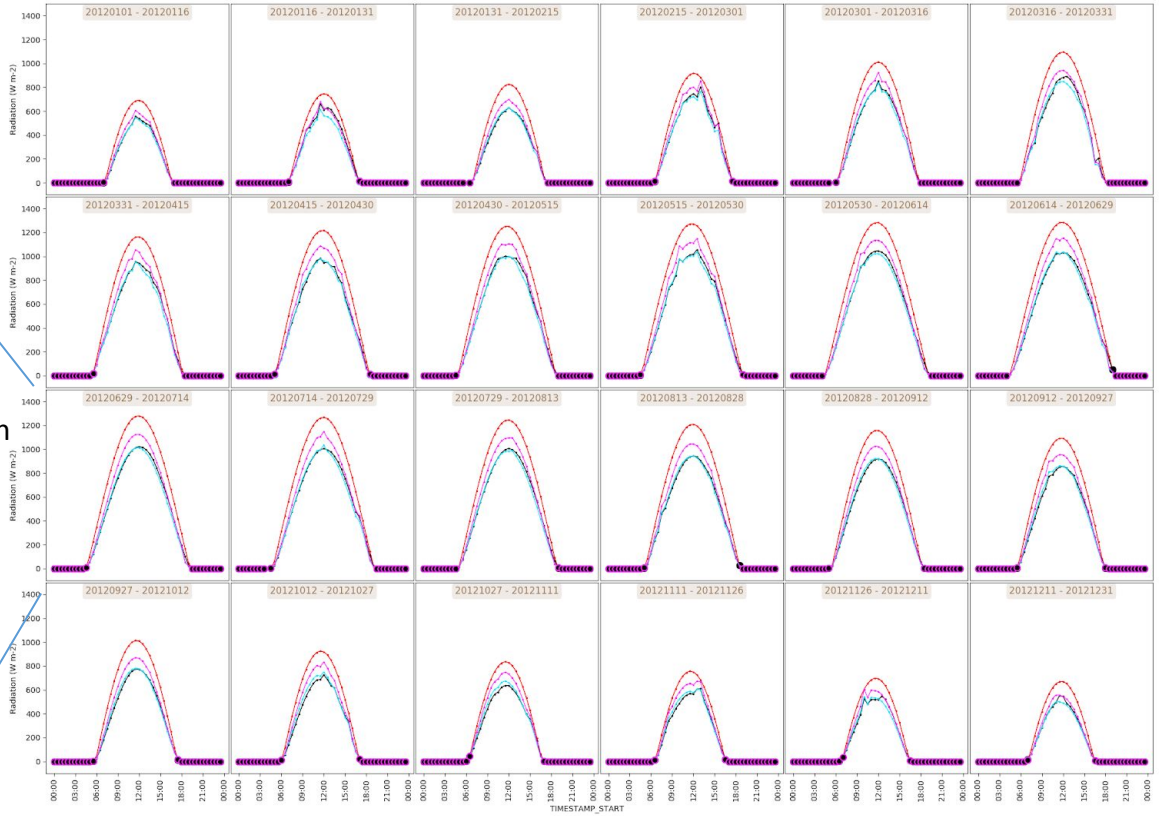
Timestamp Alignment Check

- Misspecified timestamps
- Time zone specification
- Daylight saving
- Datastream synchronization

SW_1_1_1 has max cross-correlation 0.997 at lag 0 (0% exceed SW_IN_POT)
PPFD_IN_1_1_1 has max cross-correlation 0.998 at lag 0 (0% exceed SW_IN_POT)

Max Diurnal Composite - 15 day window

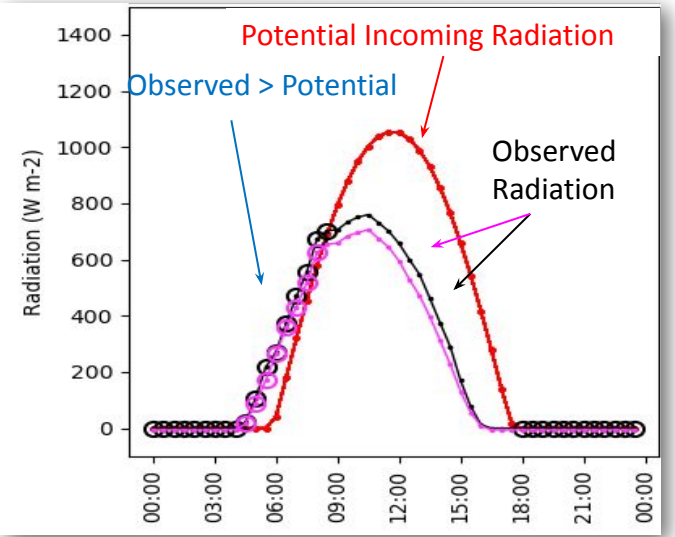




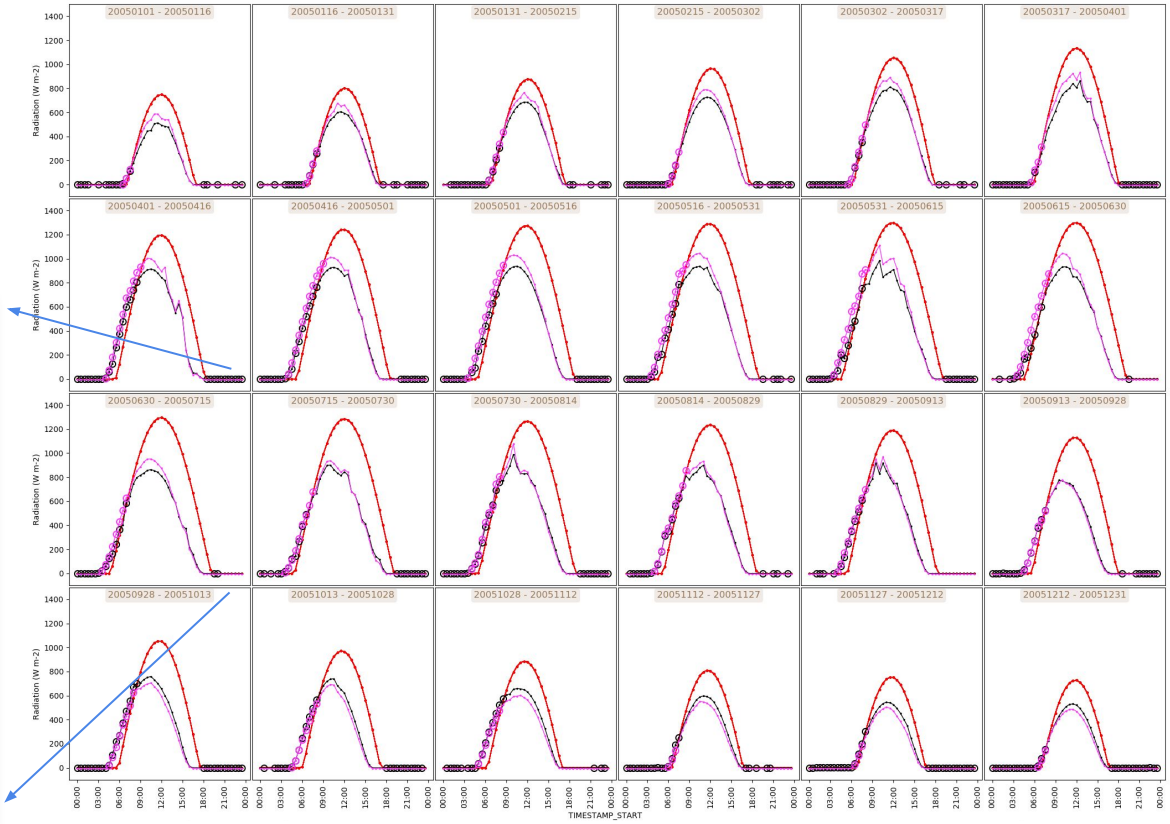
Timestamp Alignment Check

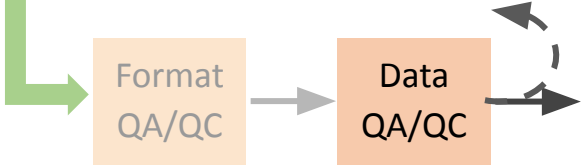
- Consistent shift
  - Timestamp – End or Start
  - Time zone specification
- Inconsistent shift
  - Daylight saving
  - Clock resetting

Max Diurnal Composite - 15 day window



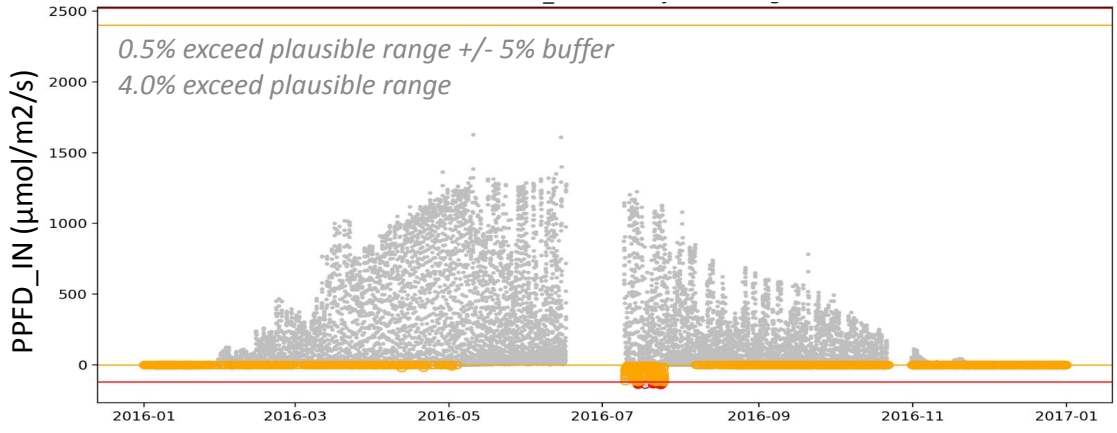
SW_1_1_1 has max cross-correlation 0.984 at lag 2 (30% exceed SW_IN_POT)
PPFD_IN_1_1_1 has max cross-correlation 0.978 at lag 2 (28% exceed SW_IN_POT)

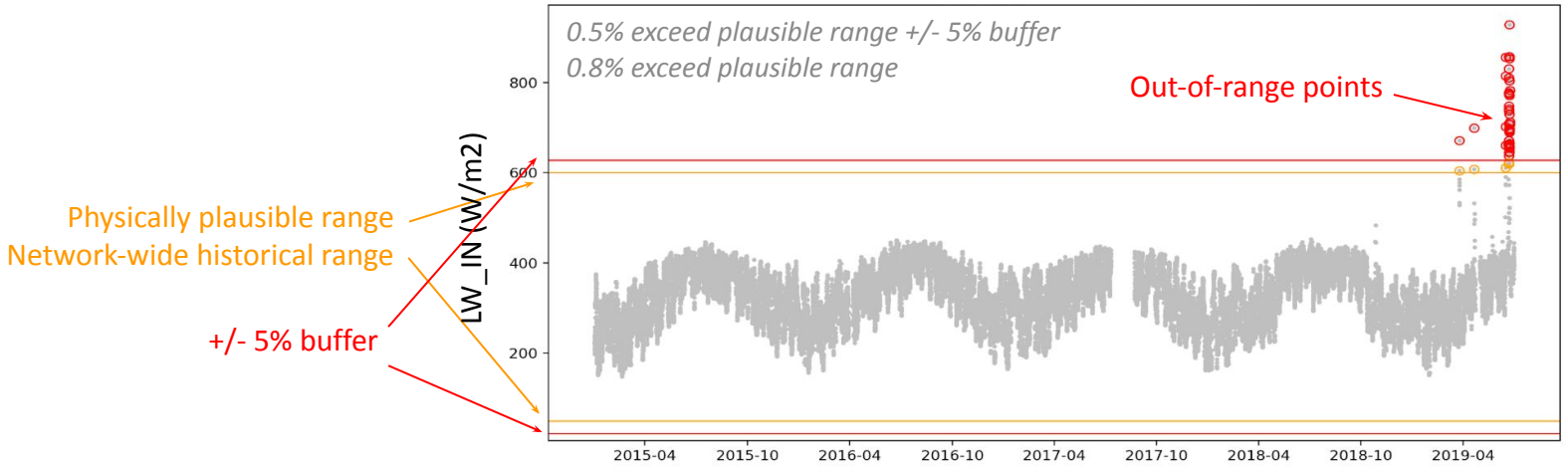




Physical Range

- Outlier / Erroneous data
- Wrong units
- Misspecified missing values





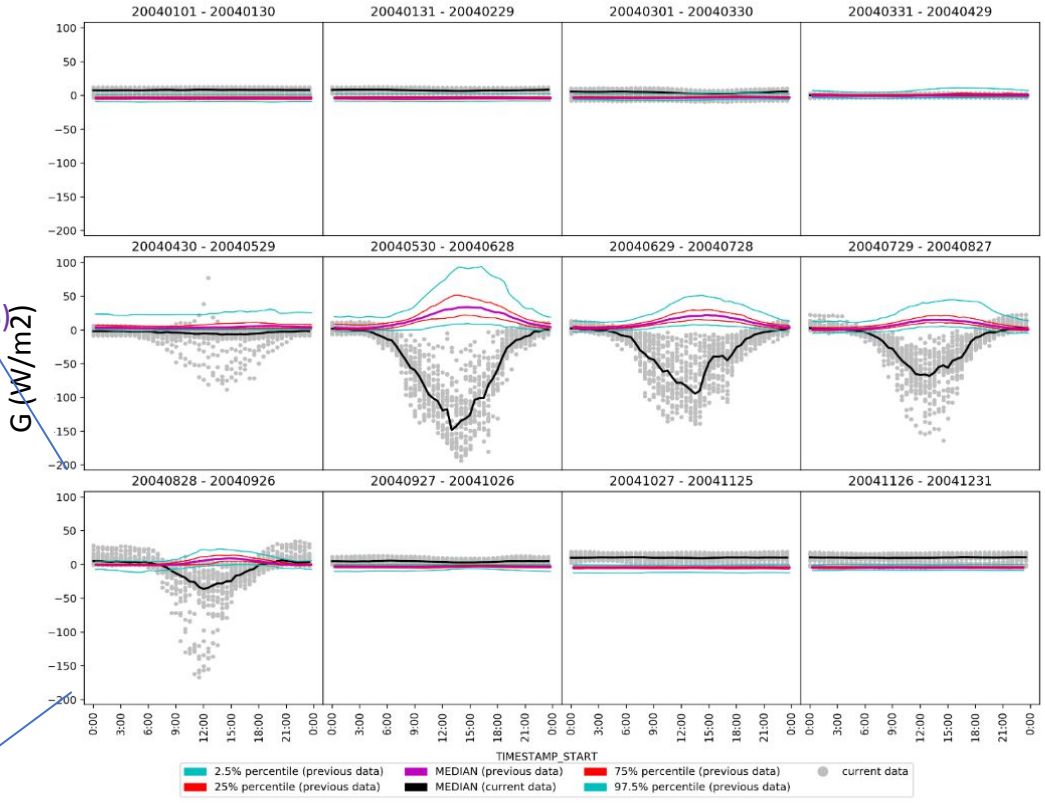
Format QA/QC → Data QA/QC

Diurnal-Seasonal Pattern

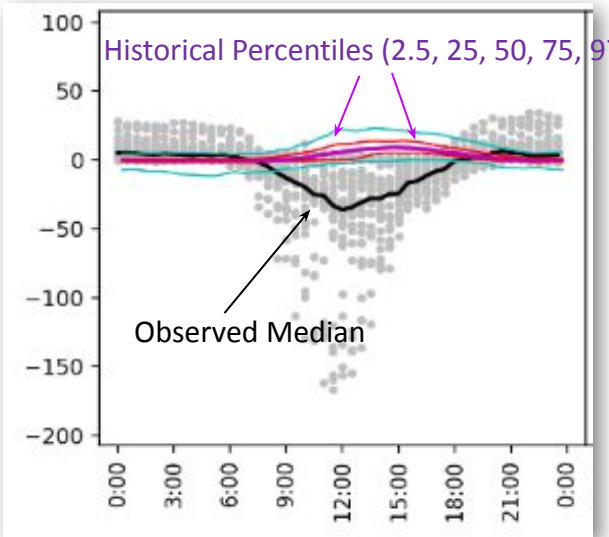
Most Variables Returning Sites

- Physically unlikely values
- Outlier
- Sign convention

40% observed data within historical 95% range
25% observed data within historical Interquartile (25-75%) range



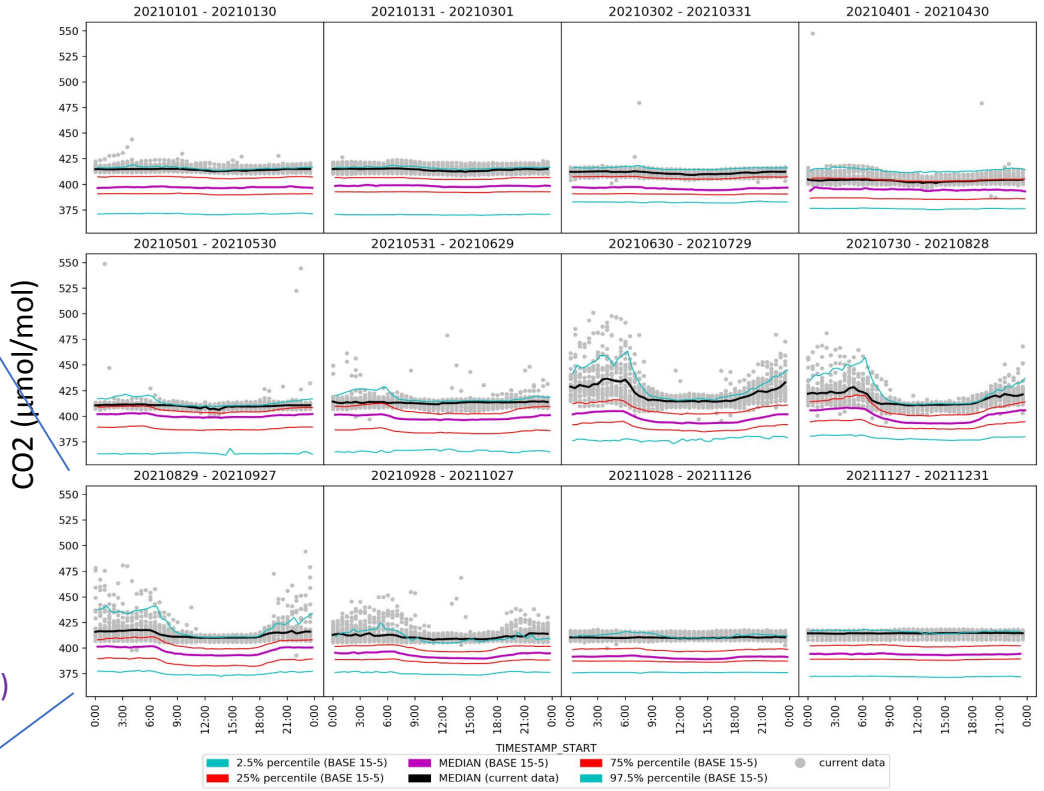
Mean Diurnal Composite - 1 month window



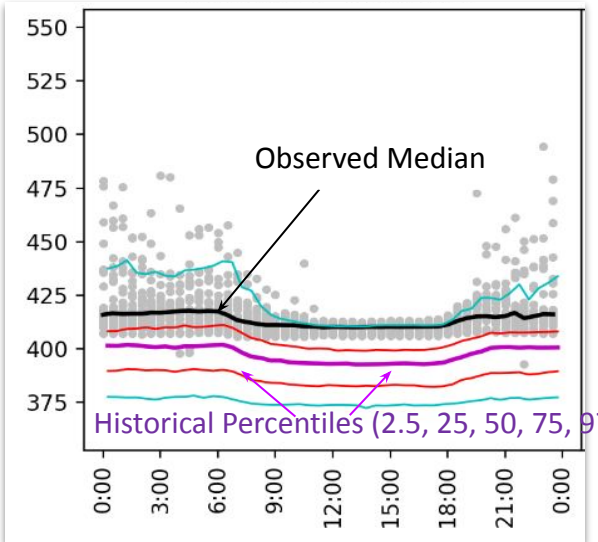
Diurnal-Seasonal Pattern

- Shifted full range
- Change of variability

78% observed data within historical 95% range
8% observed data within historical Interquartile (25-75%) range



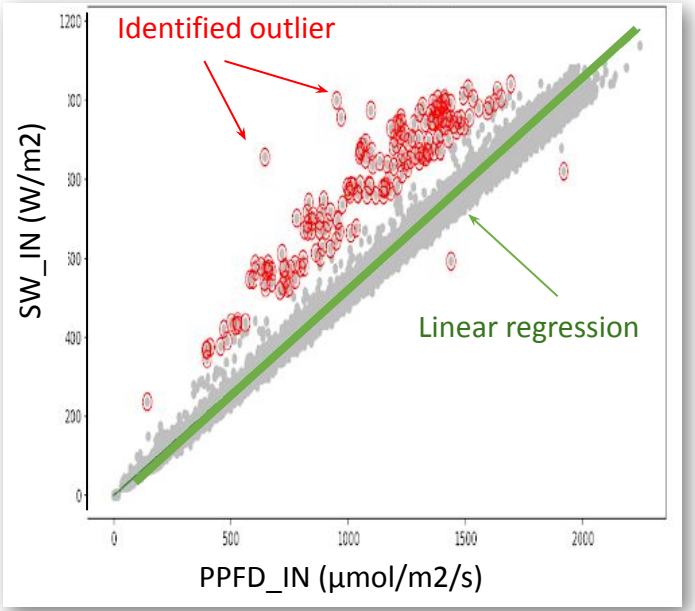
Mean Diurnal Composite - 1 month window

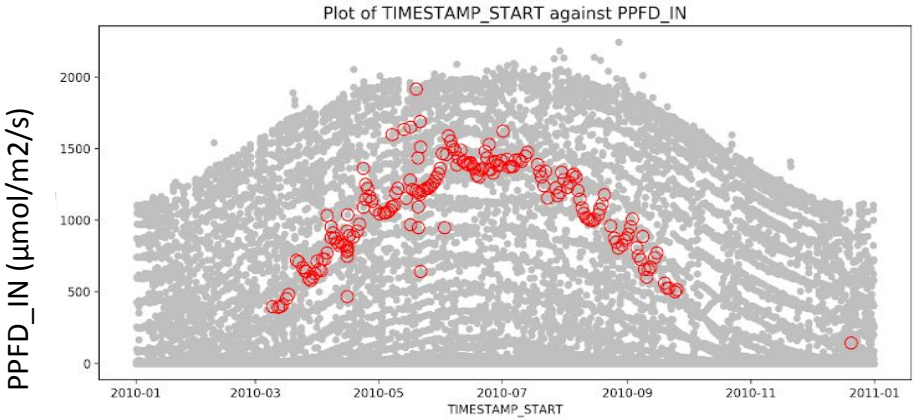


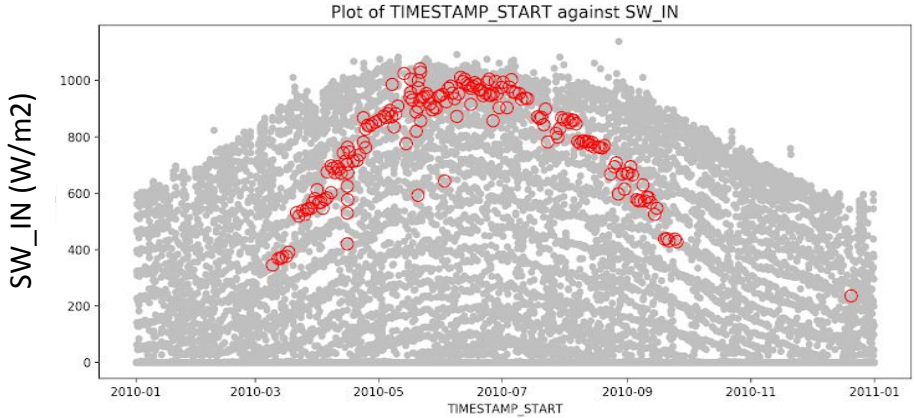
Multivariate Comparison

- Short-term Inconsistency
  - Sensor malfunction
  - Shaded radiation
  - Contamination

2% deviated from linear relationship
Linear regression slope = 0.52, R2 = 0.95



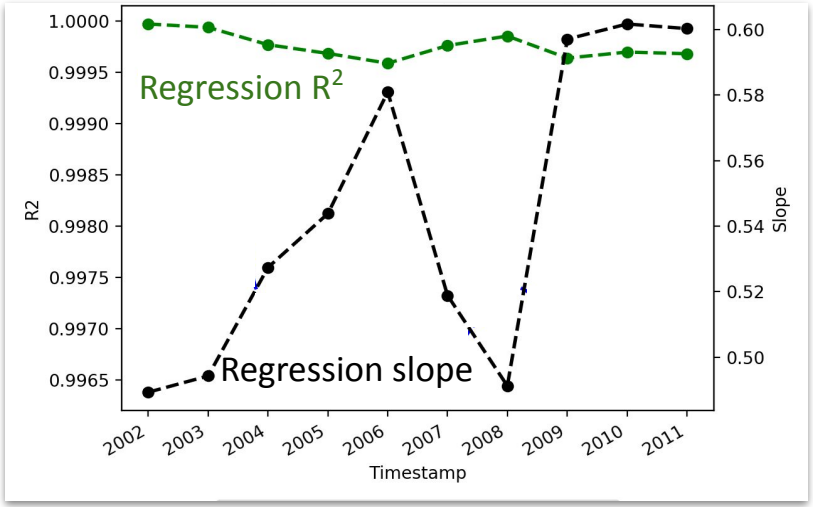


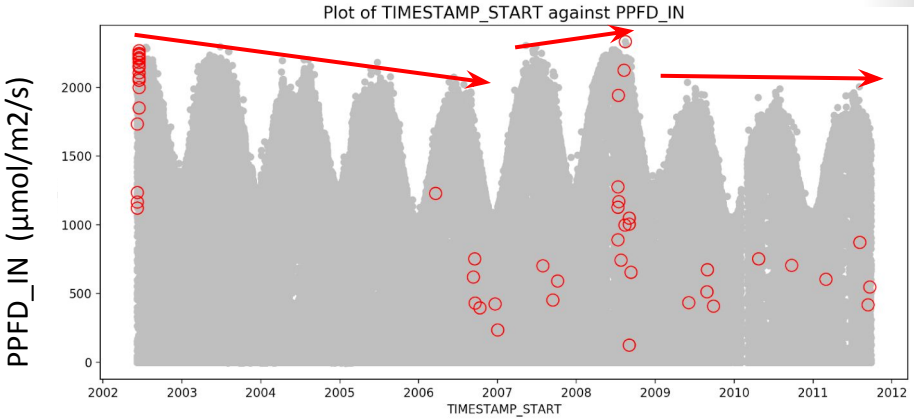


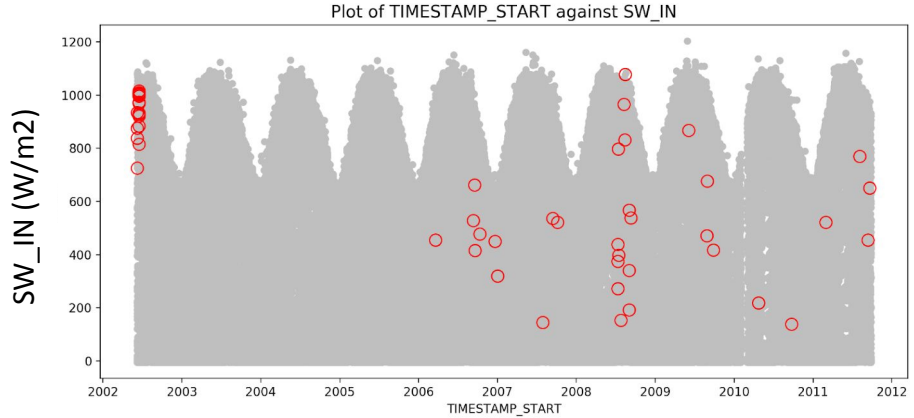
Multivariate Comparison

- Long-term trend or step change
  - Sensor Degradation
  - Replacement of sensor
  - Change of measurement location

Max changes of linear regression slope: 25%







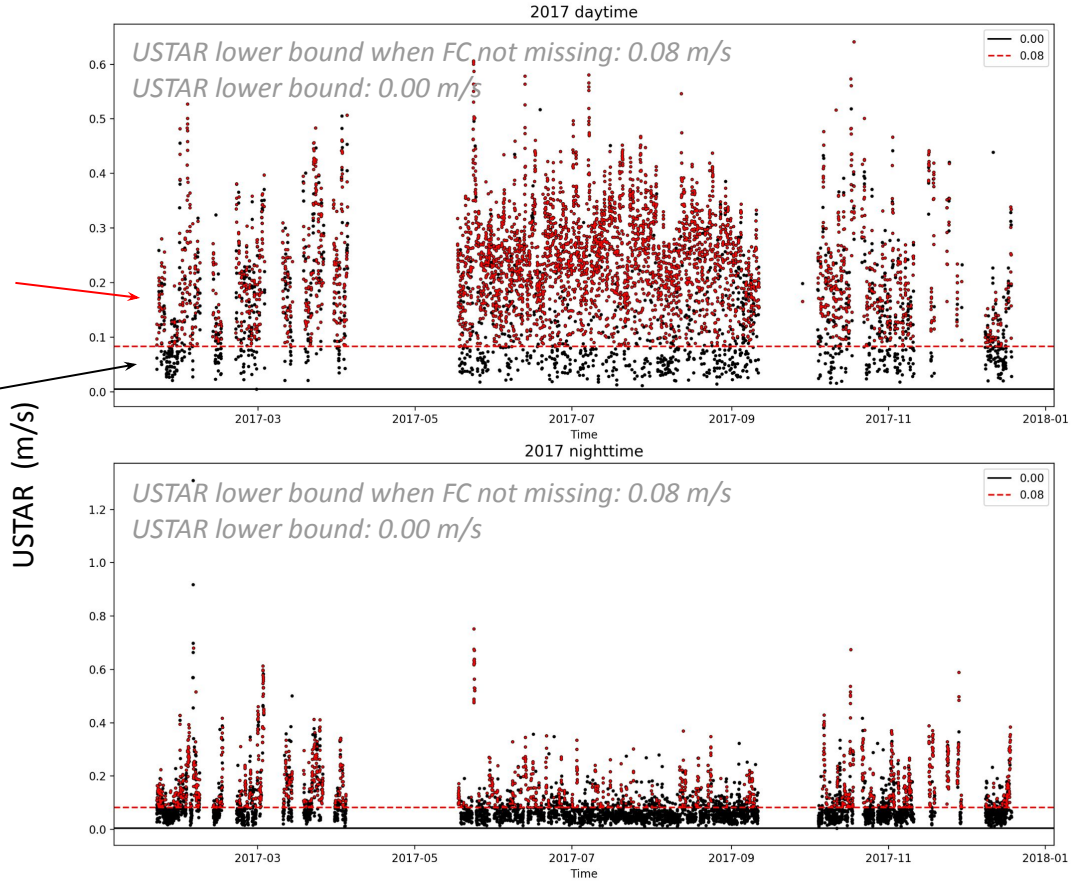
Format QA/QC → Data QA/QC

FC-USTAR Filtering

- Filtered USTAR
- Filtered FC by USTAR

USTAR when FC not missing

USTAR when FC missing

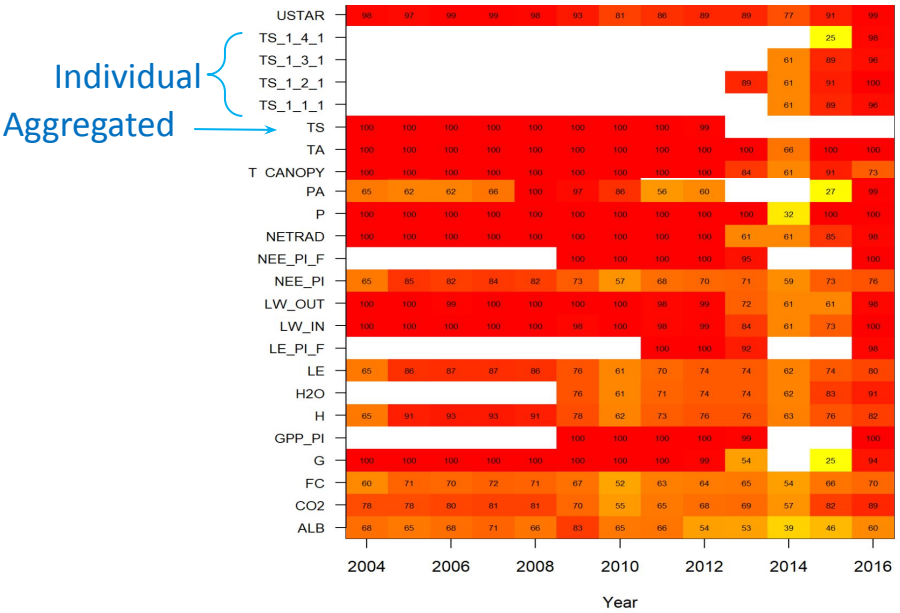


Format QA/QC → Data QA/QC

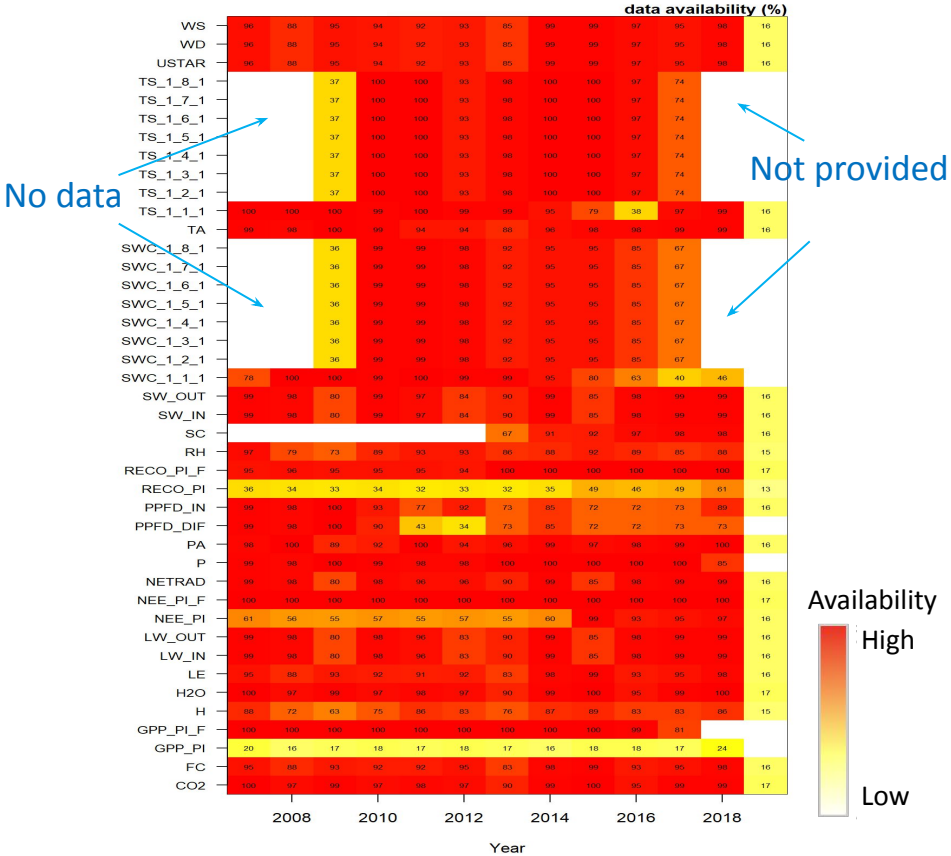
Variable Availability

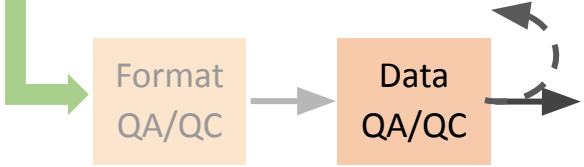
- Long gaps
- Missing mandatory variables
- Inconsistent variable naming / qualifier

All empty variable-years: TS_1_1_1, TS_1_2_1, TS_1_3_1 (2004-2013), TS (2013-2016).....



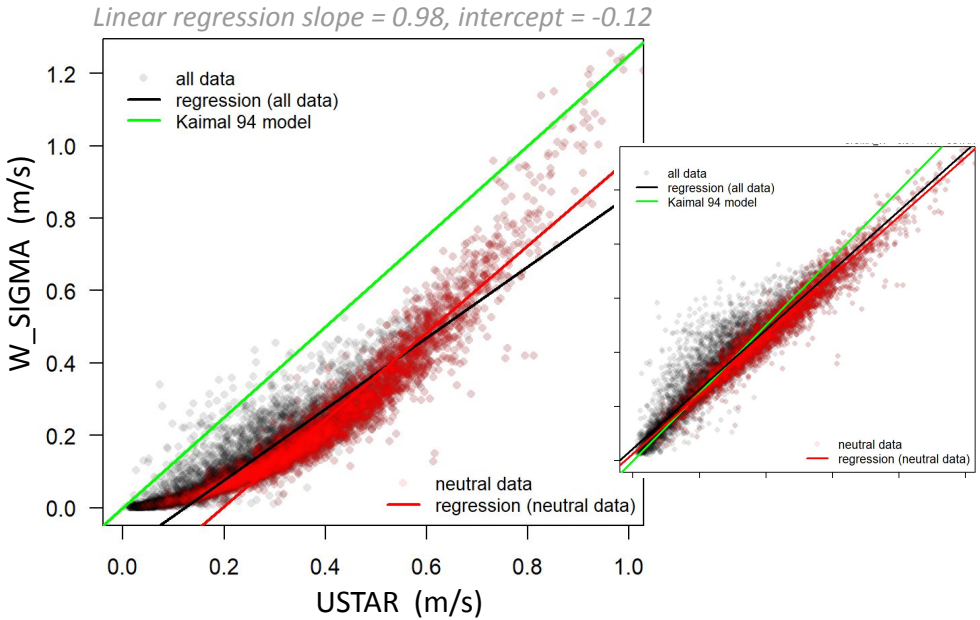
All empty variable-years: TS_1_1_1 (2007-2008, 2018-2019), SWC_1_1_1 (2007-2008, 2017-2018).....

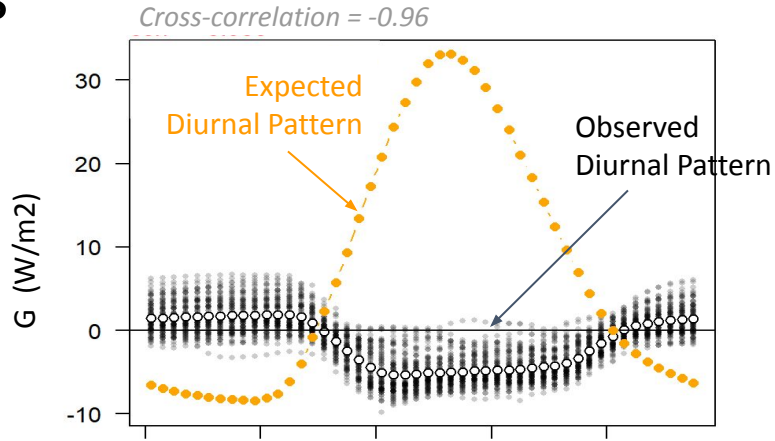


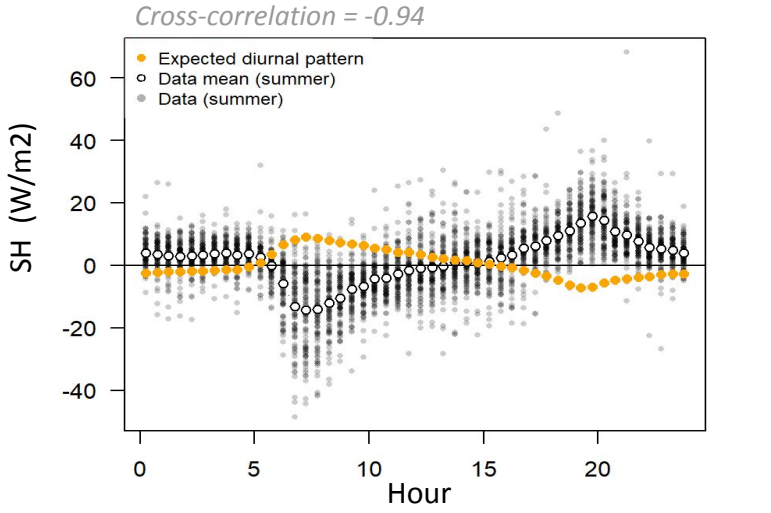


Planned Checks

- SIGMA family (standard deviation)
- Sign convention check
- Variability check (excessive, dampened)
- Unit check (ratio-percentage, metric conversion)



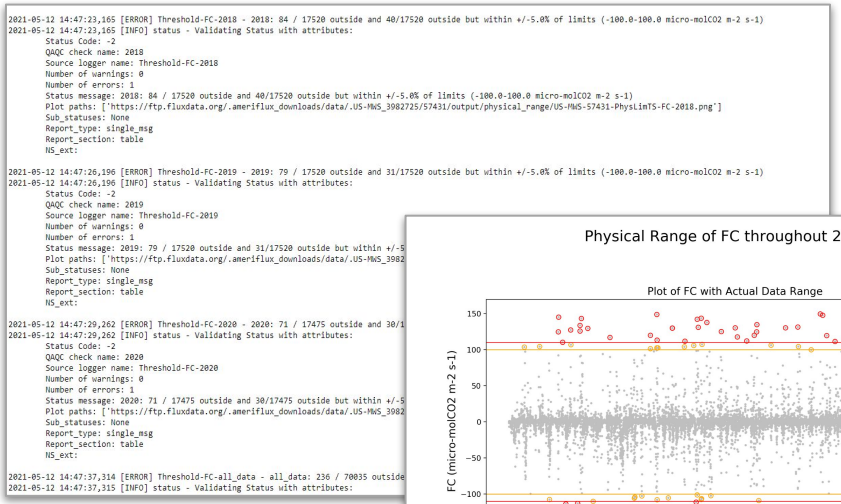


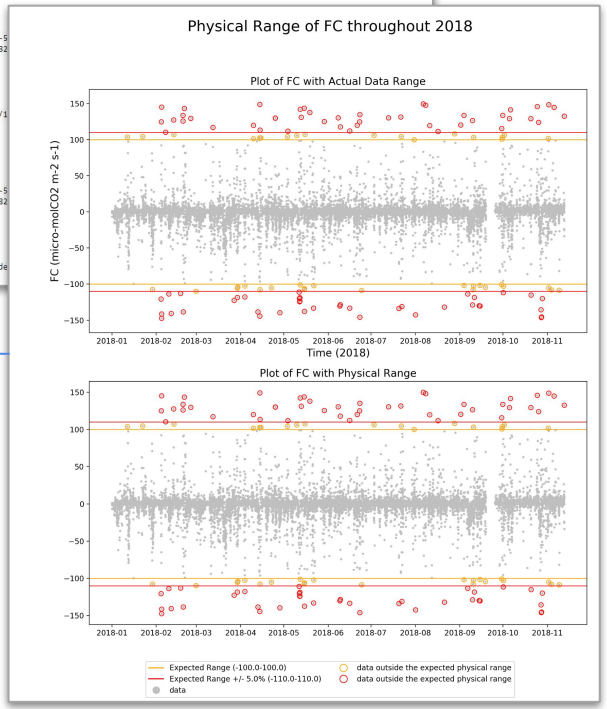


Format QA/QC → Data QA/QC

Data QA/QC Output

- Public FTP
  - logs/
  - output figures/
    - diurnal_seasonal/
    - multivariate_intercomparison/
    - physical_range/
    - timeshift/
    - ustar_filter_test/
  - Intermediate files/





- Parent Directory
- US-MWS-57431-PhysLimTS-CO2-2017.png
- US-MWS-57431-PhysLimTS-CO2-2018.png
- US-MWS-57431-PhysLimTS-CO2-2019.png
- US-MWS-57431-PhysLimTS-CO2-2020.png
- US-MWS-57431-PhysLimTS-CO2-all_data.png
- US-MWS-57431-PhysLimTS-CO2_MIXING_RATIO-2017.png
- US-MWS-57431-PhysLimTS-CO2_MIXING_RATIO-2018.png
- US-MWS-57431-PhysLimTS-CO2_MIXING_RATIO-2019.png
- US-MWS-57431-PhysLimTS-CO2_MIXING_RATIO-2020.png
- US-MWS-57431-PhysLimTS-CO2_MIXING_RATIO-all_data.png
- US-MWS-57431-PhysLimTS-FC-2017.png
- US-MWS-57431-PhysLimTS-FC-2018.png
- US-MWS-57431-PhysLimTS-FC-2019.png
- US-MWS-57431-PhysLimTS-FC-2020.png
- US-MWS-57431-PhysLimTS-FC-all_data.png
- US-MWS-57431-PhysLimTS-FC_SSITC_TEST-2017.png
- US-MWS-57431-PhysLimTS-FC_SSITC_TEST-2018.png
- US-MWS-57431-PhysLimTS-FC_SSITC_TEST-2019.png
- US-MWS-57431-PhysLimTS-FC_SSITC_TEST-2020.png
- US-MWS-57431-PhysLimTS-FC_SSITC_TEST-all_data.png
- US-MWS-57431-PhysLimTS-FETCH_70-2017.png

AMP Review

Format QA/QC → Data QA/QC

Data QA/QC and Report Email

[AmeriFlux] (QAQC-2405) Data Results | CC-XXX HH 20020101 - 20180803 | Using uploads through Sep 19, 2018

Inbox ×

AMF Data Team (AMF-JIRA) via berkeley.edu
to dschristianson

AMF Data Team **commented** on QAQC-2405

Re: Data Results | CC-XXX HH 20020101 - 20180803 | Using uploads through Sep 19, 2018

Dear Tower Team

Thank you for your data submissions for CC-XXX site.

In the context of the new processing for AmeriFlux data products, we are applying a new Data QA/QC scheme that follows the independent analysis of your data and help identify potential issues in data formats and contents earlier in the pipeline.

Briefly, Data QA/QC includes the inspection of sign conventions, ranges, diurnal-seasonal patterns, and potential outliers of var SW_IN) are also analyzed to detect potentially erroneous data. The comparison of measured radiation (e.g., PPFD_IN, SW_IN) t for a given location (i.e., SW_IN_POT) is also analyzed to check the timestamp specification and alignment.

In analyzing your data, we have the following questions where we request your expert opinion and suggestion. Please note that some i your site. Please verify, clarify, or correct the following issues before we can make your data available as an AmeriFlux BASE data prod please upload files using https://ameriflux.lbl.gov/data/upload-data/.

[Data QA/QC]

- Issue #1 .....
- Issue #2 .....
- Issue #3 .....

We hope that this will not take too much time from your work, but it will help to make your data more robust and clear. You can view th https://ameriflux.lbl.gov/qaqc-reports-data-team/.

Please reply to this email with any questions. You can track communications on this Data QA/QC at QAQC-2405 using your AmeriFlux account l

Best regards and thanks for the collaboration,
AmeriFlux Data Team

FTP link to Data QA/QC, where you can access all figures and intermediate files generated during Data QA/QC:
ftp://ftp.fluxdata.org/.ameriflux_downloads/data/CC-XXX/######/####/output

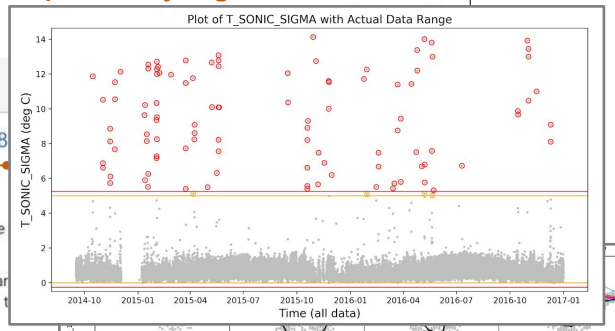
Format QA/QC reports associated with this Data QA/QC, where you can glance at the file sources used in this Data QA/QC:
http://ameriflux.lbl.gov/qaqc-report/?site_id=CC-XXX&report_id=######

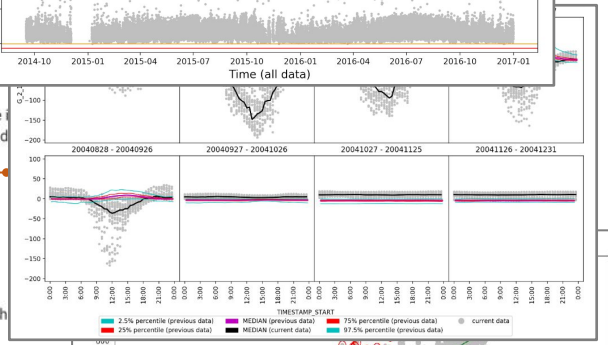
Data QA/QC Summary

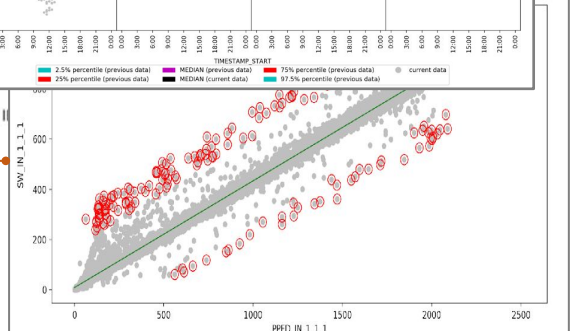
AMP-Summarized Data Issues

Additional Links

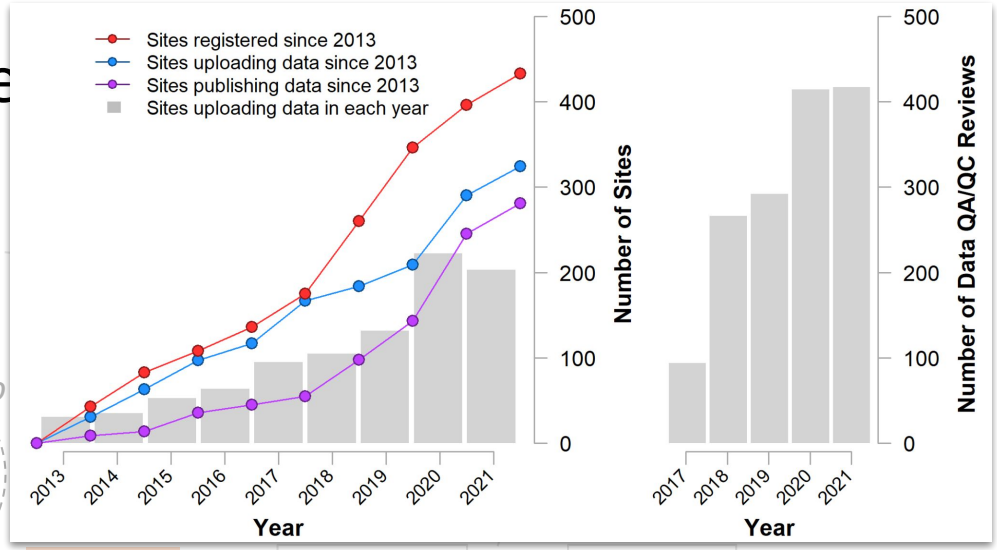
Explanatory Figures

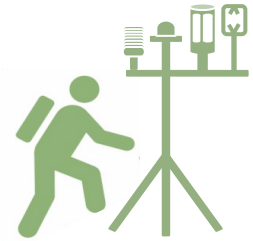






AmeriFlux Flux/Met data Pipe





Site Teams

Submit Flux/Met data

~ 1-6 months

~ 1-2 days

Iteration to

per upload
~ Immediately

batch
~ 1-2 months

Format QA/QC → Data QA/QC → BASE Publish → ONEFlux Processing → FLUXNET Publish

AmeriFlux Management Project (AMP)

BASE Data Product

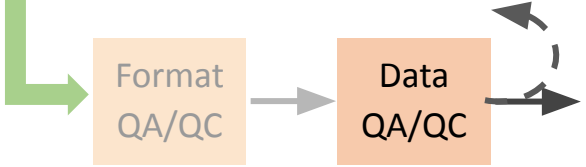
- Provided as submitted by site team
- QA/QC on general data quality
- All FP variables supported
- All levels of aggregation supported

Available to Data Users

FLUXNET Data Product

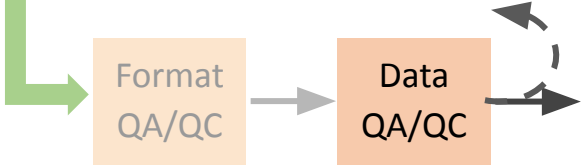
- Value-added data product
- QA/QC on ONEFlux requirements
- Subset of standard FP variables
- Site-representative aggregation

Question?



Data QA/QC Next Stage

- Consideration
  - Keep up with growth of networks & submissions
  - Develop while maintaining data service
  - Enable self-review & quick assessment
  - Gain users' feedback on automatic QA/QC
  - Refine check modules & rules for passing/warning
  - Scalable workflow for adding new check modules



Data QA/QC Next Stage

- Consideration
  - Keep up with growth of networks & submissions
  - Develop while maintaining data service
  - Enable self-review & quick assessment
  - Gain users' feedback on automatic QA/QC
  - Refine check modules & rules for passing/warning
  - Scalable workflow for adding new check modules

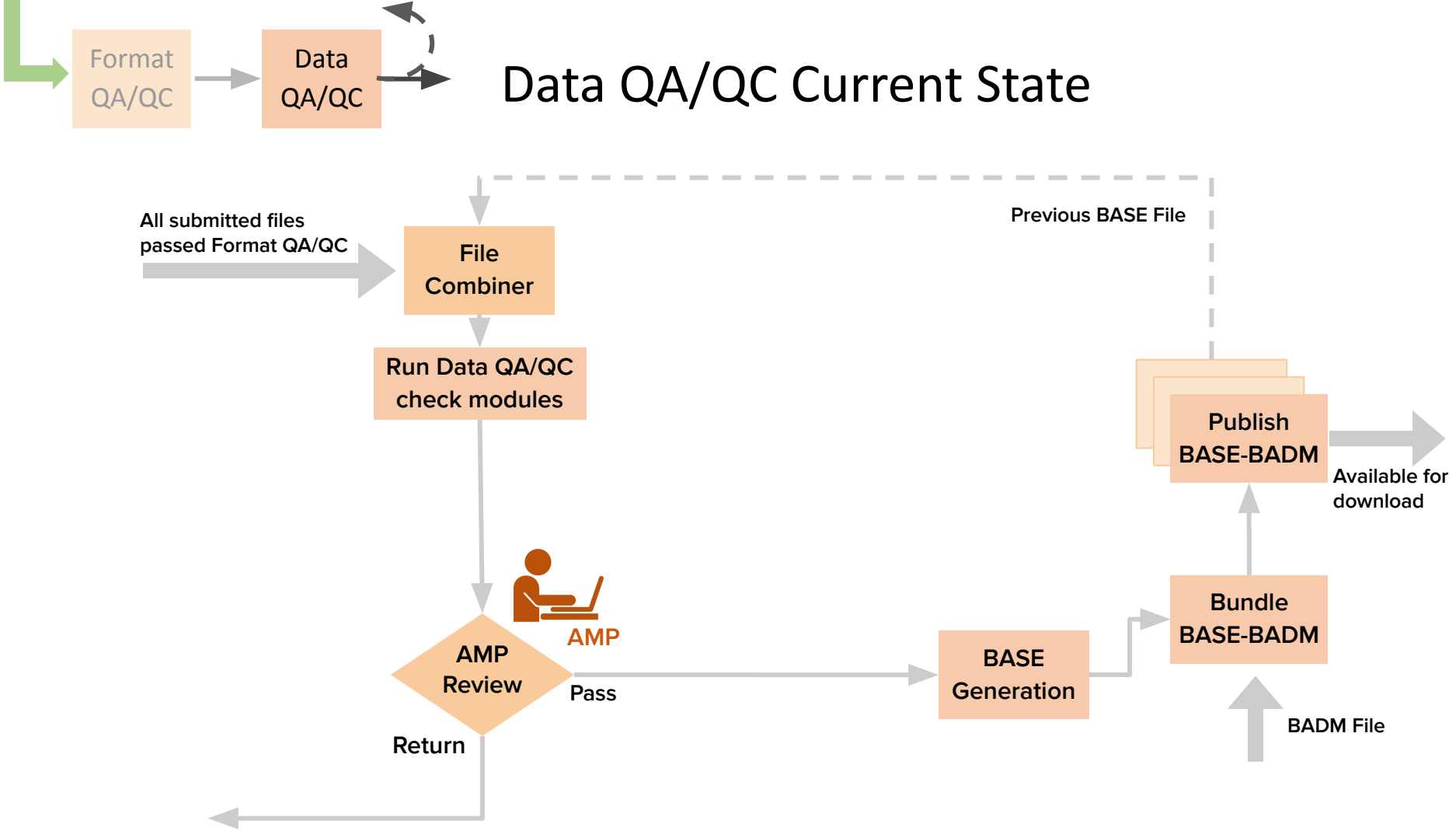
Phase II: Automatic data QA/QC report
- Most returning sites
- Self-explanatory QA/QC report
- Develop and implement new check modules

Phase I: Enable self-review for trained users
- Core, NEON, volunteer returning sites (limited capacity)
- Training + Summary statistics + Linked figures
- Gain users' feedback designing Data QA/QC reports
- Re-evaluate and refine check modules

Current state

2022/Jan | Apr | 2022/Jul | Oct | 2023/Jan | 2023/Apr | Jul

Data QA/QC Current State



Format QA/QC → Data QA/QC

Data QA/QC Next Stage (Phase I)

All submitted files passed Format QA/QC

File Combiner

Previous BASE File

1 Training Webinar

2 Run Data QA/QC check modules + New summary statistic table(s) generated

3 Send Report

No

Yes

4

Site Teams

Self Review

Pass

Return

AMP

AMP Review

Pass

Return

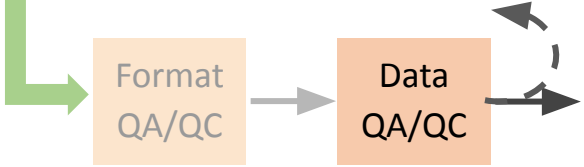
BASE Generation

Bundle BASE-BADM

BADM File

Publish BASE-BADM

Available for download



Data QA/QC Next Stage (Phase I)

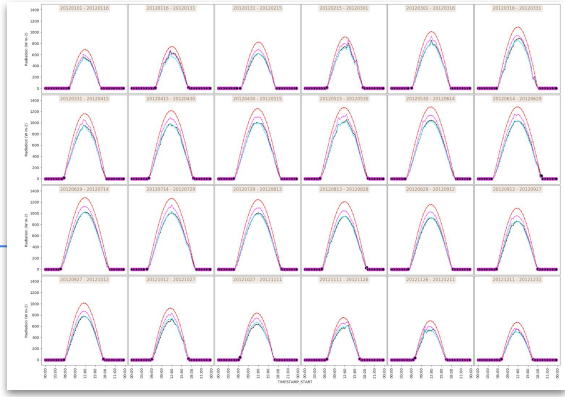
- Public FTP
  - logs/
  - output figures/
  - Intermediate files/

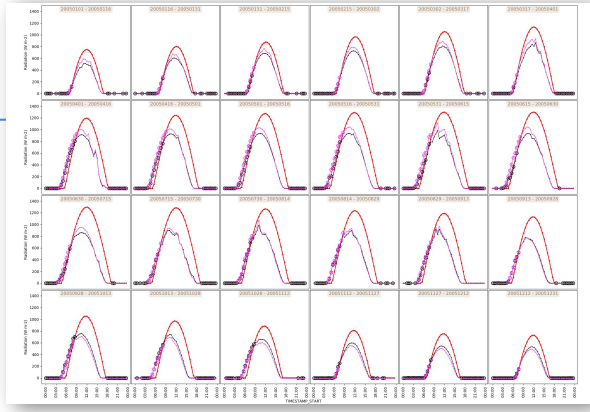
- Summary statistics/
  - Timestamp alignment/
  - ……

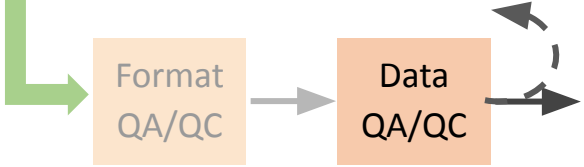
Year	max(abs(R_{xy}))	t_{max}	P_{day} (%)	P_{night} (%)	Figure link
2016	0.98	0	5	5	<url>
2017	0.97	0	1	2	<url>
2018	0.99	0	0	1	<url>
2019	0.95	2	10	10	<url>

- **max(abs(R_{xy}))** : maximum absolute cross correlation (R_{xy}) between the time series X and Y.
- **t_{max}** : the timestep shift at which the max(abs(R_{xy})) is found. t_{max} equating zero indicates that time series X and Y are aligned.
- **P_{day}** : percentage of occasions that measured radiation exceed potential incoming radiation in daytime.
- **P_{night}** : percentage of occasions that measured radiation exceed potential incoming radiation in nighttime.

- Parent Directory
- US-NR1_timestamp_shift_radiation_1998.png
- US-NR1_timestamp_shift_radiation_1999.png
- US-NR1_timestamp_shift_radiation_2000.png
- US-NR1_timestamp_shift_radiation_2001.png
- US-NR1_timestamp_shift_radiation_2002.png
- US-NR1_timestamp_shift_radiation_2003.png
- US-NR1_timestamp_shift_radiation_2004.png
- US-NR1_timestamp_shift_radiation_2005.png
- US-NR1_timestamp_shift_radiation_2006.png
- US-NR1_timestamp_shift_radiation_2007.png
- US-NR1_timestamp_shift_radiation_2008.png
- US-NR1_timestamp_shift_radiation_2009.png
- US-NR1_timestamp_shift_radiation_2010.png
- US-NR1_timestamp_shift_radiation_2011.png
- US-NR1_timestamp_shift_radiation_2012.png
- US-NR1_timestamp_shift_radiation_2013.png
- US-NR1_timestamp_shift_radiation_2014.png
- US-NR1_timestamp_shift_radiation_2015.png
- US-NR1_timestamp_shift_radiation_2016.png
- US-NR1_timestamp_shift_radiation_2017.png
- US-NR1_timestamp_shift_radiation_2018.png
- US-NR1_timestamp_shift_radiation_2019.png
- US-NR1_timestamp_shift_radiation_2020.png
- US-NR1_timestamp_shift_radiation_2021.png







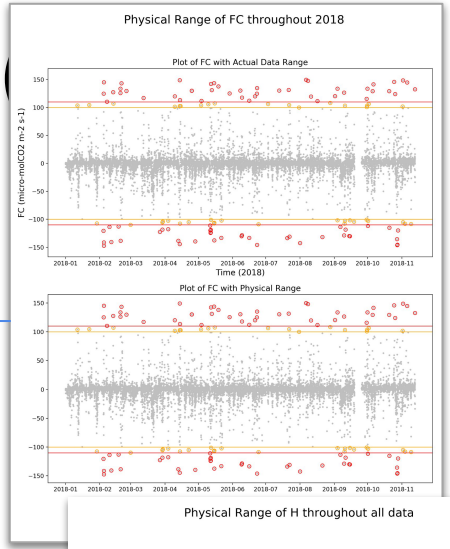
Data QA/QC Next Stage (

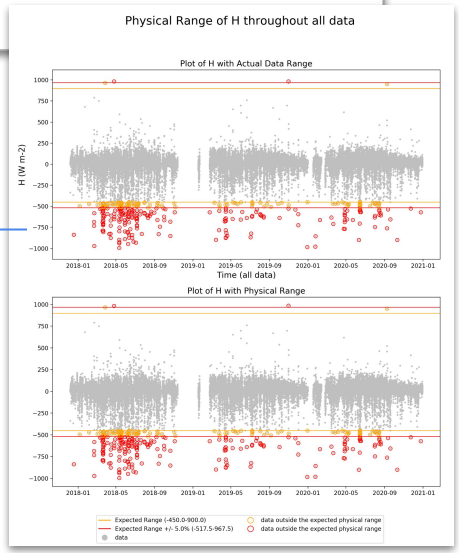
- Public FTP
  - logs/
  - output figures/
  - Intermediate files/

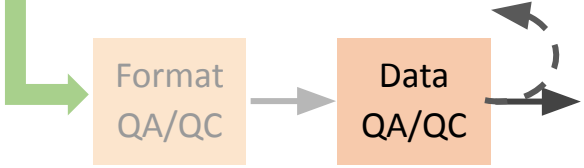
- Summary statistics/
  - Timestamp alignment/
  - Physical range/
  - ……

Variable	Year	P_{soft_flag} (%)	P_{hard_flag} (%)	Figure link
FC	2018	2.5	1.5	<url>
FC	2019	0.1	0	<url>
LE	2018	0	0	<url>
LE	2019	0	0	<url>
H	2018	8.9	5.6	<url>
H	2019	0	0	<url>

- P_{soft_flag} : percentage of data points outside the expected physical range, but within the buffer range (+/-5% of physical range)
- P_{hard_flag} : percentage of data points outside the expected physical range plus the buffer range.







Data QA/QC Next Stage (Phase I)

[AmeriFlux] (QAQC-2405) Data Results | CC-XXX HH 20020101 - 20180803 | Using uploads through Sep 19, 2018

Inbox ×

AMF Data Team (AMF-JIRA) via berkeley.edu
to dschristianson

Thu, Sep 20, 12:37 PM

AMF Data Team **commented** on QAQC-2405

Re: Data Results | CC-XXX HH 20020101 - 20180803 | Using uploads through Sep 19, 2018

Dear Tower Team

Thank you for your data submissions for CC-XXX site.

Data QA/QC Summary

In the context of the new processing for AmeriFlux data products, we are applying a new Data QA/QC scheme that follows the Format QA/QC. We believe that these checks can provide an independent analysis of your data and help identify potential issues in data formats and contents earlier in the pipeline.

Briefly, Data QA/QC includes the inspection of sign conventions, ranges, diurnal-seasonal patterns, and potential outliers of variables. Multivariate relations (e.g., WS vs USTAR, PPFD_IN vs SW_IN) are also analyzed to detect potentially erroneous data. The comparison of measured radiation (e.g., PPFD_IN, SW_IN) to the maximum, top of the atmosphere radiation expected for a given location (i.e., SW_IN_POT) is also analyzed to check the timestamp specification and alignment.

In analyzing your data, we have the following questions where we request your expert opinion and suggestion. Please note that some issues we identify could be normal and expected at your site. Please verify, clarify, or correct the following issues before we can make your data available as an AmeriFlux BASE data product. If you decide to resubmit a corrected version, please upload files using https://ameriflux.lbl.gov/data/upload-data/.

[Data QA/QC]

Links to Summary Statistics Tables

- Link to Summary Statistics (Timestamp alignment)
- Link to Summary Statistics (Physical range)
- Link to Summary Statistics (Multivariate comparison)
- ……

We hope that this will not take too much time from your work, but it will help to make your data more robust and clear. You can view the status of all of your uploaded files at https://ameriflux.lbl.gov/qaqc-reports-data-team/.

Please reply to this email with any questions. You can track communications on this Data QA/QC at QAQC-2405 using your AmeriFlux account ID and password to login.

Best regards and thanks for the collaboration,
AmeriFlux Data Team

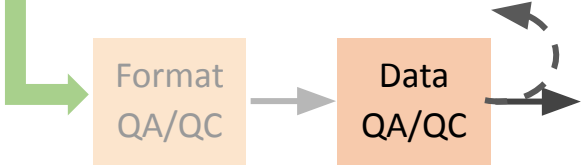
Links to Training Materials

- Link to Training Materials (<URL>)

Additional Links

FTP link to Data QA/QC, where you can access all figures and intermediate files generated during Data QA/QC:
ftp://ftp.fluxdata.org/.ameriflux_downloads/data/CC-XXX/######/####/output

Format QA/QC reports associated with this Data QA/QC, where you can glance at the file sources used in this Data QA/QC:
http://ameriflux.lbl.gov/qaqc-report/?site_id=CC-XXX&report_id=######



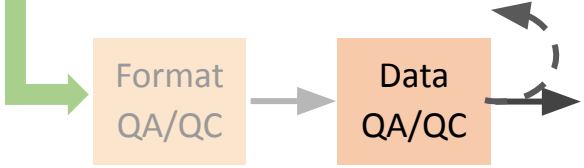
Data QA/QC Next Stage (Phase I)

Detailed Timeline 2022

- Now - March: Seek for participating sites (invitation sent later)
- June / July (TBD): Training webinar
- June - September:
  - Submit data as planned & self-review using summary statistics (interim)
  - Participate in the individual user interview
- September (TBD): Re-evaluation webinar

What to expect?

- Participate in the training and re-evaluation webinars
- Submit data (new quarters/year) at least once & self-review using summary statistics
- Participate in the individual user interview (1-2 hour feedback session, 1-2 times)



Data QA/QC Future (Phase II, draft plan)

All submitted files passed Format QA/QC

File Combiner

Previous BASE File

1

Run Data QA/QC check modules

Rule for passing each Data QA/QC module

2

Self-explanatory report

Send Auto Report

No

AMP

AMP Review

Pass

Return

Yes

3

Self Review

Pass

Return

Site Teams

BASE Generation

Bundle BASE-BADM

Publish BASE-BADM

Available for download

BADM File

Questions / Feedback

- How do you think about the new self-review process?
- Would the summary statistics be manageable? Or anything else?
- Which part of the data pipeline or QA/QC can be improved?
- Would you be interested in participating in the self-review?
- Anything you'd like to see in future data pipeline?
- .....

AMERIFLUX MANAGEMENT PROJECT

Related links

- https://ameriflux.lbl.gov/data/data-processing-pipelines/
- https://ameriflux.lbl.gov/half-hourly-hourly-data-upload-format/
- https://ameriflux.lbl.gov/community/amp-webinar-series/

ameriflux-support@lbl.gov

@AmeriFlux